BIOCHEMISTRY

Accelerated Publications

X-ray Structure of Cerulean GFP: A Tryptophan-Based Chromophore Useful for Fluorescence Lifetime Imaging[†,‡]

Gabrielle D. Malo,[#] Lauren J. Pouwels,[#] Meitian Wang,[§] Andrzej Weichsel,[⊥] William R. Montfort,[⊥] Mark A. Rizzo,[||] David W. Piston,[○] and Rebekka M. Wachter*[,#]

Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604, Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, Arizona 85721, and Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennesee 37232

ABSTRACT: The crystal structure of the cyan-fluorescent Cerulean green fluorescent protein (GFP), a variant of enhanced cyan fluorescent protein (ECFP), has been determined to 2.0 Å. Cerulean bears an internal fluorophore composed of an indole moiety derived from Y66W, conjugated to the GFP-like imidazolinone ring via a methylene bridge. Cerulean undergoes highly efficient fluorescence resonance energy transfer (FRET) to yellow acceptor molecules and exhibits significantly reduced excited-state heterogeneity. This feature was rationally engineered in ECFP by substituting His148 with an aspartic acid [Rizzo et al. (2004) *Nat. Biotechnol. 22*, 445], rendering Cerulean useful for fluorescence lifetime imaging microscopy (FLIM). The X-ray structure is consistent with a single conformation of the chromophore and surrounding residues and may therefore provide a structural rationale for the previously described monoexponential fluorescence decay. Unexpectedly, the carboxyl group of H148D is found in a buried position, directly contacting the indole nitrogen of the chromophore via a bifurcated hydrogen bond. Compared to the similarly constructed ECFP chromophore, the indole group of Cerulean is rotated around the methylene bridge to adopt a cis-coplanar conformation with respect to the imidazolinone ring, resulting in a close edge-to-edge contact of the two ring systems. The double-humped absorbance spectrum persists in single-crystal absorbance measurements, casting doubt on the idea that ground state conformational heterogeneity forms the basis of the two overlapping transitions. At low pH, a blue shift in absorbance of 10−15 nm suggests a pH-induced structural transition that proceeds with a time constant of 47 (±2) min and is reversible. Possible interpretations in terms of chromophore isomerization are presented.

To be useful as biocatalysts, therapeutic compounds, or biosensors, wild-type proteins often require substantial improvements in their biochemical and biophysical properties. Characteristics to be optimized include protein stability (*1*), substrate specificity (*2*), or optical properties essential in live cell imaging applications (*3*). Two experimental approaches are routinely combined to confer desirable properties to proteins: rational design, involving site-specific amino acid replacements based on structural or functional arguments, and directed evolution, involving the creation of

genetic libraries followed by selection for key characteristics (2). Both methods have been successfully applied to members of the family of green fluorescent protein (GFP)[1]-like proteins, a set of genetically encodable fluorescent markers used as biotechnological tools in live cell imaging and biosensor applications (4, 5). A number of recent examples testify to the remarkable achievements that can be attained when applying structure-guided directed evolution to fluorescent proteins (6–8). Live-cell Förster resonance energy transfer (FRET) experiments using genetically encoded fluorophores provide dynamic insight into molecular interactions within a distance of about 5 nm (9). Several examples illustrate the striking development of improved GFP variants for use as imaging tools in FRET measurements, such as the rational design of yellow fluorescent proteins (YFPs) (10, 11), which include Citrine (12), a variant that has proven particularly useful as an acceptor fluorophore in combination with cyan-fluorescent donor GFPs.

In all known naturally occurring GFP-like proteins, the chromophore's conjugated π-system is derived from three amino acids, where the central residue is always a tyrosine (Tyr66 in wild-type GFP) (13, 14). Wild-type cyan-fluorescent GFPs isolated from reef-building corals exhibit emission wavelengths around 485 nm, suitable for FRET experiments in combination with Citrine (15, 16). However, reef GFPs naturally exist as tight tetramers, limiting their usefulness as reporter genes because of their tendency toward intracellular aggregation (4). Extensive genetic engineering efforts have led to the development of a monomeric teal-fluorescent protein with peak emission at 492 nm (7) and a number of improved blue-fluorescent proteins (17). In spite of these efforts, the engineering of nonaggregating GFPs that emit cyan light from a tyrosine-based chromophore has to our knowledge not been achieved yet.

Some years ago, GFP-based cyan fluorescent proteins were constructed by replacing Tyr66 with a tryptophan residue, yielding a modified chromophore structure in which the bulk of the π system comprises an indole moiety (emission ~ 480 nm) (18). Several cyan clones carrying the Y66W mutation are currently used as FRET donors in combination with the yellow Citrine acceptor: enhanced cyan fluorescent protein (ECFP) (19), Cerulean (20) and related variants (21), and CyPet (22). However, the broad spectra of these variants lead to cross-excitation due to overlap of the absorbance bands of the cyan and yellow fluorophores. Therefore, GFP-based FRET systems continue to be limited by their small dynamic range (3–5-fold) and poor signal-to-noise ratio (20). To alleviate these problems, fluorescence lifetime imaging microscopy (FLIM) has recently been promoted as an intensity-independent method to monitor FRET signals in biological samples. This method takes advantage of a shortening of the donor lifetime when energy transfer to the acceptor occurs. However, FLIM is only useful if the donor excited-state decay is essentially monoexponential.

Cerulean GFP is an optimized FRET donor molecule that was rationally designed by Piston and co-workers to eliminate the excited-state heterogeneity of ECFP (20). In ECFP, a single-exponential fit to the fluorescence lifetime is not feasible due to multiple components, whereas Cerulean exhibits essentially homogeneous excited-state decay kinetics, rendering this protein useful for lifetime imaging. The original design strategy to generate Cerulean was based on the ECFP X-ray structure (23), which suggested two alternate conformations for Tyr145 and His148, with either one or the other side chain solvent-exposed. However, the ECFP chromophore itself adopted a single conformation only, that of a coplanar π-system that places the bulk of the indole group in trans to the five-membered imidazolinone ring (23). To remove the conformational heterogeneity of residues 145 and 148, the H148D mutation was introduced into ECFP, with the intent of stabilizing solvent-exposure of residue 148 by substituting the imidazole with a carboxyl group (20). Remarkably, this substitution had indeed the desired effect of an essentially homogeneous fluorescence lifetime. Other optical properties, such as brightness and quantum yield, were subsequently improved by the introduction of the S72A (24) and Y145A substitutions, yielding Cerulean GFP (ECFP/S72A/Y145A/H148D) (20).

Live cell imaging experiments have demonstrated that Cerulean GFP undergoes highly efficient FRET to yellow acceptor molecules (25, 26). Additionally, Cerulean exhibits more than twice the brightness of ECFP and CyPet, emitting light at fluorescence intensities similar to Citrine (20). The reduced excited-state heterogeneity leads to a remarkable improvement in the signal-to-noise ratio of FRET intensity measurements (20), in part because homo-FRET resulting from multiple excited states has effectively been eliminated. Surprisingly, the absorption and fluorescence spectra of Cerulean remain similar to those of ECFP over a broad pH range, with two overlapping bands observed for both excitation and emission spectra (20). Here, we report that the high-resolution X-ray structure of Cerulean GFP is consistent with a single ground state conformation for the chromophore and its protein environment, even though the broad absorbance doublet persists in the crystal.

MATERIALS AND METHODS

Protein Expression, Purification, Crystallization, and Data Collection. 6His-tagged Cerulean GFP was expressed in *E. coli* M15 using the pQE-9 expression vector (Qiagen). Liquid cultures were induced at OD600 = 1.5 with 1 mM IPTG for 5.5 h at 37 °C and purified using standard Ni-NTA (Qiagen) affinity purification procedures (10). Rod-shaped crystals were grown by hanging drop vapor diffusion at 4 °C. Drops consisted of 2 μL of protein added to 2 μL of mother liquor. Mother liquor contained 100 mM sodium acetate, 19% v/v PEG 4000, and 1 mM EDTA. The pH of the mother liquor was measured to be 5.0, utilizing a pH electrode after 2-fold dilution of the mother liquor with water. Crystals were transferred to a cryoprotectant solution (mother liquor plus 20% v/v glycerol) and stored at 77 °K in liquid nitrogen until data collection. Diffraction data were collected at 100 K using a 3 × 3 CCD array (ADSCQ315) detector at

[#] Arizona State University.

[⊥] University of Arizona.

[○] Vanderbilt University.

[‖] Current address: Department of Physiology, University of Maryland School of Medicine, Baltimore, MD 21201.

[§] Current address: Swiss Light Source, Paul Scherrer Institut, Villigen PSI, Switzerland 5232.

[1] Abbreviations: FLIM, fluorescence lifetime imaging microscopy; FRET, Förster resonance energy transfer (or fluorescence resonance energy transfer); GFP, green fluorescent protein; ECFP, enhanced cyan fluorescent protein; YFP, yellow fluorescent protein.

Table 1: Crystallographic and Refinement Statistics

Data Collection and Processing	
space group	$P2_12_12_1$
unit cell dimensions	$a = 51.232$ $b = 63.339$ $c = 69.821$, $\alpha = \beta = \gamma = 90°$
detector	3×3 CCD array (ADSCQ315)
wavelength (Å)	1.0000
resolution (Å)	20−2.00
high-resolution shell (Å)	2.11 − 2.00
total observations	90,354
unique reflections	15,852
redundancy	5.7 (5.8)[a]
Av I/σ	5.1 (1.8)[a]
completeness (%)	99.5 (99.4)[a]
R_{merge} (%)[b]	8.2 (20.7)
Refinement	
resolution range (Å)	19.92−2.00
number reflections	15,023
R_{cryst} (%)[c]	18.8
R_{free} (%)[c]	26.2
rms deviation bond lengths (Å)	0.017
rms deviation bond angles (°)	1.866
average B factors, all atoms (Å^2)	13.24
total atoms	1,941
model	residues 1−229, 176 waters

[a] Values within parentheses refer to the high-resolution shell (2.11−2.00 Å). [b] $R_{merge} = \sum |I_{hkl} − <I>|/\sum<I>$, where $<I>$ is the average of individual measurements of I_{hkl}. [c] R_{cryst} and $R_{free} = \sum_h (||F(h)_{obs}| − |F(h)_{calc}||)/|F(h)_{obs}|$ for reflections in the working and test sets (5% of all data).

the Advanced Light Source Beamline 5.0.2 (Lawrence Berkeley National Laboratory).

Structure Determination and Refinement. Indexing and integration were carried out using MOSFLM (*27*), and the data were scaled and merged with SCALA in the *CCP4* suite of programs (*28*). Data were processed in a $P2_12_12_1$ cell (a = 51.232 Å, b = 63.339 Å, c = 69.821 Å, $\alpha = \beta = \gamma = 90°$) (Table 1). Molecular replacement was carried out using the program MOLREP (*29*) using ECFP in the major conformation A′ as a search model (PDB ID code 1OXD) (*23*). After 5% of the data was removed to determine R_{free}, rigid body and positional refinement was carried out using REFMAC5 (*30*) from 46.9 to 2.0 Å resolution. Model building of the protein (without chromophore) was carried out using Xfit (*31*), making use of $2|F_o| − |F_c|$ and $|F_o| − |F_c|$ maps. After several rounds of refinement, R_{cryst} and R_{free} were 26.2 and 31.1%, respectively. The α-carbon positions of the two published ECFP structures (PDB ID code 1OXD for the major A′, and 1OXE for the minor B′ conformation (*23*)) were superimposed onto the Cerulean model, and the chromophore was modeled using coordinates from the aligned ECFP structure in the minor conformation. Water molecules were modeled into 3σ positive difference density peaks where suitable hydrogen-bonding partners were available. Adjustment of the model according to strong positive and negative difference density resulted in placement of the H148D side chain close to the chromophore indole group. The Cerulean chromophore was best modeled in the cis conformation, whereas both published ECFP X-ray structures contain trans chromophores. Additional refinement was carried out in the absence of chromophore torsion restraints and with relaxed planarity restraints, to allow for potential twisting of the π system. After density modification and several rounds of refinement, R_{cryst} fell to 18.8% (R_{free} = 26.2%). The final model consisted of residues 1−229 and

151 waters. Electron density was not observed for the 18-residue 6HisTag, the carboxy terminal residues 230−238, and several solvent-exposed side chains.

Absorbance and Fluorescence Measurements. All absorbance scans were collected on a Shimadzu UV-2401PC UV−vis spectrophotometer at ambient temperature. Buffers contained 100 mM NaCl, 1 mM EDTA, and 50 mM of one of the following: sodium citrate pH 2.5, sodium acetate pH 5.0, piperazine-1,4-bis(2-ethanesulfonic acid) (PIPES) pH 6.0, *N*-cyclohexyl-2-aminoethanesulfonic acid (HEPES) pH 7.0, HEPES pH 8.0, *N*-cyclohexyl-3-aminopropanesulfonic (CAPS) pH 9.0, or CAPS pH 10.0. Absorbance measurements of protein samples equilibrated in these buffers were collected at a concentration of 1.0 mg/mL. To monitor any pH-dependent absorbance change with time, protein stored at pH 8.0 was rapidly diluted into pH 5.0, pH 6.0 buffer, or pH 7.0 buffer (or vice versa), to a final protein concentration of 0.4 mg/mL. Subsequently, absorbance scans were collected every 10 min for the first 90 min, then every half hour up to 150 min. The absorbance at the wavelength of maximal change (439 nm) was normalized by dividing this value by the optical density at 452 nm (the isosbestic point), and the change in normalized absorbance y as a function of time t was fit to a unimolecular rate equation of the form $y = y_{max} − [e^{−kt}y_{max}]$, where k is the observed rate constant.

Fluorescence emission scans were collected at ambient temperature at a protein concentration of 0.1 mg/mL, using a Jobin Ivon Horiba Fluoromax-3 fluorimeter, with integration time set to 0.1 s, and excitation and emission slit widths set to 2 nm band-pass. For measurements carried out at pH 5.0, the excitation and emission wavelengths were as follows: $\lambda_{ex} = 400$ nm, $\lambda_{em} = 420−650$ nm; $\lambda_{ex} = 424$ nm, $\lambda_{em} = 435−650$ nm; $\lambda_{ex} = 439$ nm, $\lambda_{em} = 450−650$ nm; $\lambda_{ex} = 455$ nm, $\lambda_{em} = 465−650$ nm. For measurements carried out at pH 8.0, the excitation and emission wavelengths were as follows: $\lambda_{ex} = 407$ nm, $\lambda_{em} = 430−650$ nm; $\lambda_{ex} = 435$ nm, $\lambda_{em} = 445−650$ nm; $\lambda_{ex} = 454$ nm, $\lambda_{em} = 465−650$ nm; $\lambda_{ex} = 468$, $\lambda_{em} = 475−650$ nm. To measure the change in emission intensity over time, measurements were repeated at specified time intervals after rapid dilution from pH 8.0 to pH 5.0.

Single-Crystal Absorbance Measurements. The absorption spectra of frozen single crystals of Cerulean GFP were measured using a micro-spectrophotometer equipped with a CCD array detector (Spectral Instruments, Inc., Tucson, AZ) and focusing optics (4DXray Systems, Uppsala, Sweden), mounted on an optical bench. The optical elements were focused to produce a 15 μm incident beam generated by a Xenon lamp light source. Crystals were mounted on cryoloops (Hampton Research), frozen in liquid nitrogen, and positioned on the spectrophotometer in a 100 K nitrogen cryostream (Oxford CryoSystems, Oxford, UK). The spectra were collected in the wavelength range 350−650 nm, with the incident beam approximately normal to the face of the platelike crystals. Meaningful spectra could only be collected on crystals thinner than 50 μm, as thicker crystals exceeded the upper limit of detection by the spectrophotometer. The crystal yielding the best spectra was a thin plate ∼60 μm in the longest dimension, grown by hanging drop vapor diffusion at 4 °C in the following mother liquor: 16% PEG 4000, 100 mM sodium acetate, and 1 mM EDTA, at pH 5.0.
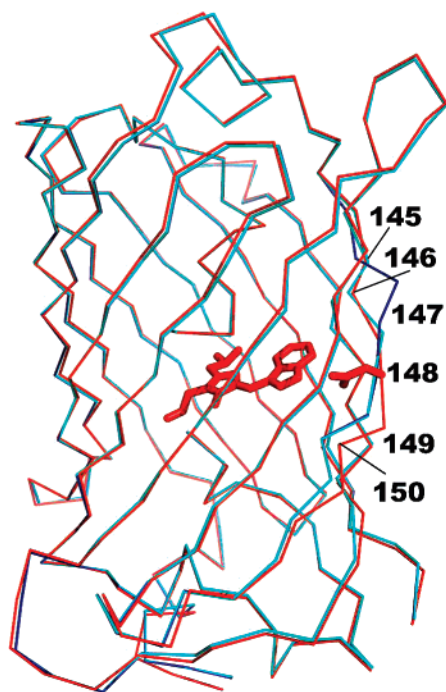
FIGURE 1: Overlay of α-carbon trace of Cerulean (red), ECFP A′ conformation (dark blue), and ECFP B′ conformation (cyan). The chromophore and the side chain of H148D are shown as stick models.

RESULTS

Description of the Cerulean Structure. Cerulean GFP (ECFP/S72A/Y145A/H148D) is a GFP variant bearing a non-native chromophore derived from the internal S65T-Y66W-Gly67 tripeptide (*20*). To discover structural features responsible for the relatively homogeneous nature of the excited state, Cerulean GFP was expressed in *Escherichia coli*, purified, and crystallized at pH 5.0 in the space group $P2_12_12_1$ ($a = 51.232$ Å, $b = 63.339$ Å, $c = 69.821$ Å). The structure was determined by molecular replacement using ECFP as the search model (PDB code 1OXD) (*23*) and refined to 2.0 Å, with one monomer per asymmetric unit and a final R-factor of 18.8% (Table 1). The overall fold is that of the 11-stranded GFP β-barrel, with the autocatalytically formed chromophore positioned within the central helix (*13, 14*). The root-mean-square deviations of the Cerulean α-carbons from those of ECFP in the A′ (major, 1OXD) and B′ (minor, 1OXE) conformation are 0.54 and 0.40 Å, respectively (Figure 1).

Surprisingly, the Cerulean structural arrangement between residues 145 and 148 closely follows that of the ECFP B′ conformation (Figure 1). In addition, the hydrophilic carboxyl group of H148D points toward the protein interior, whereas the hydrophobic methyl group of Y145A is flipped out into the solvent. These observations are opposite to the expectation that this β-strand would adopt the ECFP A′ conformation, to allow for solvent exposure of the engineered aspartic acid in position 148 (*20*). Although the residues comprising the β-bulge region 145−149 have B factors higher (∼22 Å²) than the average B-factor of the model (∼13 Å²), there are no crystallographic features suggesting alternate main chain or side chain conformations in Cerulean GFP.

The electron density omit map (Figure 2) unequivocally places the smaller side chain introduced by the H148D substitution within hydrogen-bonding distance to the chromophore indole nitrogen, in contrast to ECFP, where the His148 imidazole ring does not interact with the chromophore (Figure 3). The β-strand comprising residues 148−152 buckles out significantly when compared to the main chain traces of both ECFP conformations A′ and B′, perhaps to accommodate hydrogen bonding of H148D to the chromophore (Figure 1). The largest displacement, 1.4 Å, is observed for Asn149. The main chain movements observed for Cerulean along this β-strand are reminiscent of those previously described for deGFPs, dual-emission pH sensors that contain a tyrosine-based chromophore (*32*).

Cis Conformation and Intra-Chromophore Hydrogen Bonding. Contrary to ECFP, the indole group of the Cerulean chromophore adopts a cis conformation with respect to the imidazolinone ring, thus placing the bulky six-membered ring in close proximity to the heterocycle constructed from the main chain atoms of residues 65−67 (Figure 2). In spite of the isomerization around the Y66W-derived C_β−C_γ bond of the β-methylene bridge connecting the two ring systems, the Cerulean chromophore adopts an entirely planar conformation. Electron density omit maps generated during refinement did not indicate any twisting or bending of the chromophore. The angle formed by the bridging atoms C_α−C_β−C_γ (derived from Y66W) refines to 137°, significantly larger than the equivalent angle in ECFP (116°−118°) (*23*), but similar to that found in the tyrosine-based trans chromophores characterized in the kindling fluorescent protein (*33*). No crystallographic evidence for multiple chromophore conformations can be discerned, as further supported by a low average B factor of 10.5 Å² for all chromophore atoms.

Surprisingly, a close edge-to-edge contact between the two ring systems places one of the indole carbon atoms about 3.1 Å away from the imidazolinone ring nitrogen. This proximity is consistent with a close van der Waals contact but may also indicate a weak hydrogen bonding interaction of the type C−H···N, where the donor is an aromatic carbon atom (*34*). Interatomic C to N distances of 3.0−4.0 Å have previously been described in neutron crystal structures of organic molecules (*35*), and weak hydrogen bonding has recently been reported to occur between pyridyl nitrogen atoms and phenyl C−H groups in the stabilization of a supramolecular structure (*36*). Although the details of this interaction cannot be discerned at the present structural resolution, the lack of any bending or twisting of the chromophore ring system appears surprising, given the unusual placement of the bulky indole group in relation to the imidazole ring. The Cerulean imidazolinone ring occupies the same position as in ECFP (average atomic rmsd = 0.23 Å), and the rotated indole moiety fills nearly the same volume element as in ECFP (Figure 3), suggesting that isomerization of the tryptophan-derived chromophore may be a volume-conserving motion, as has been proposed for the wild-type chromophore of GFP (*37*).

A Bifurcated Hydrogen Bond of the H148D Carboxyl with the Chromophore. The cis conformation of the Cerulean chromophore allows for a direct interaction of the indole nitrogen with the engineered aspartic acid (Figures 2 and 4). The carboxyl oxygens of H148D are positioned nearly symmetrically about the indole nitrogen, consistent with the nitrogen acting as a hydrogen bond donor to both oxygen atoms in a bifurcated hydrogen bond. As often observed in
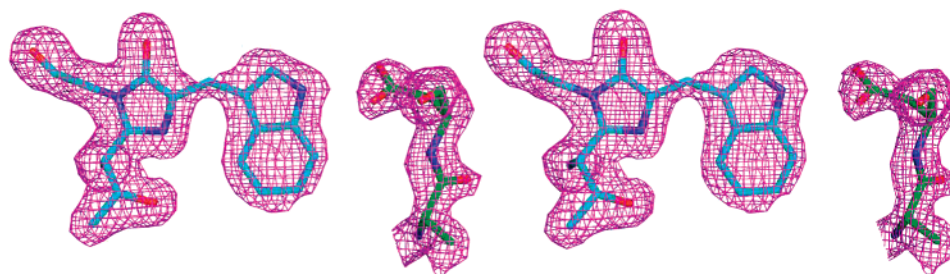
FIGURE 2: Stereoview of the $|F_o| - |F_c|$ electron density omit map, sigmaA-weighted and contoured at 3.0σ. The chromophore atoms derived from residues 65−67, and atoms of residues 147 and 148, are superimposed on the map (oxygen red, nitrogen blue, carbon green, all atoms of chromophore cyan).
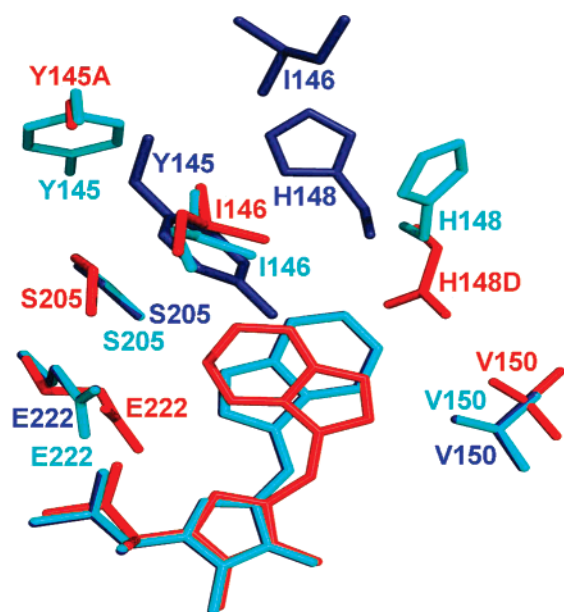

FIGURE 3: Overlays of Cerulean (red), ECFP minor (cyan), and ECFP major (dark blue). Differences in the chromophore environment are shown. Structural overlay was produced by aligning α-carbon positions.
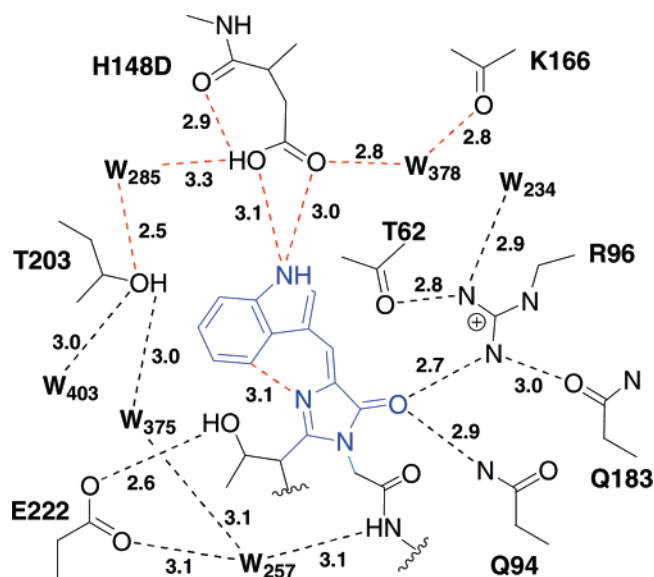

FIGURE 4: Schematic representation of Cerulean chromophore and environment illustrating likely hydrogen bonds with dashed lines (distance between heavy atoms in Å). Hydrogen bonds proposed to play a role in stabilizing the cis conformation of the chromophore are shown in red, and crystallographically ordered solvent molecules are indicated by W.

protein X-ray structures (*38*), the interaction geometry suggests that one oxygen atom of the three-center hydrogen bond, here OD2 of Asp148, may be a somewhat weaker acceptor than the other (3.1 and 3.0 Å separation, respectively, see Figure 4). H148D OD2 is found within hydrogen-bonding distance to the backbone carbonyl oxygen of the same residue, implying that OD2 bears a proton, i.e., the aspartic acid side chain is uncharged. The side chain protonation state is inferred by its proper positioning for hydrogen bonding to the main chain carbonyl group, which cannot serve as H-bond donor. Additionally, both oxygen atoms of H148D form hydrogen bonds to nearby water molecules, which in turn are hydrogen bonded to Thr203 and Lys166 (Figure 4, red dashed lines). The hydroxyl group of Thr203 is part of an extended hydrogen-bonded network that includes several ordered water molecules and extends to the Glu222 side chain, the hydroxyl group derived from S65T, and the Val68 main chain nitrogen. Although there is no structural evidence for multiple conformations, the side chain atoms of Asp148 are best modeled with slightly elevated thermal factors ($\sim$22 Å^2) compared to those of Glu222 ($\sim$14 Å^2).

Positional Adjustments of Residues in the Vicinity of the Chromophore. Despite the rotated indole group and the buried position of the H148D side chain, the Cerulean chromophore retains most of the hydrophobic contacts present in the B′ conformation of ECFP. These include contacts to the residues Phe46, Val61, Ile146, Val150, Phe165, Ile167, and Leu220 (Figure 3). In Cerulean, as in the minor conformation of ECFP, Ile146 is positioned adjacent to the chromophore indole moiety, contrary to the ECFP A′ conformation, where the equivalent volume element is occupied by Tyr145. In both ECFP conformations, the Ser205 hydroxyl forms a hydrogen bond with the indole nitrogen of the chromophore. In Cerulean, however, the side chain of Ser205 is rotated such that the hydroxyl group is shifted away by 1.3 Å, allowing space for the bulky six-membered ring of the chromophore. In addition, the side chain of Val150 has moved by 1.3 Å to make room for the indole five-membered ring, which occupies space close to Val150 in the cis conformation of the chromophore.

Optical Properties of Cerulean in Solution and in the Crystalline State. The absorbance spectrum of Cerulean is essentially invariant between pH 7.0 and 10.0, and exhibits two maxima at 435 and 454 nm, with the dominant band observed at 435 nm (Figure 5A). However, if the pH is lowered from 8.0 to 5.0, a slow spectral blue-shift of 11 and 15 nm is observed over the course of 150 min (Figure 5B), with absorbance maxima developing at 424 and 439 nm. The process is entirely reversible over a similar time period.
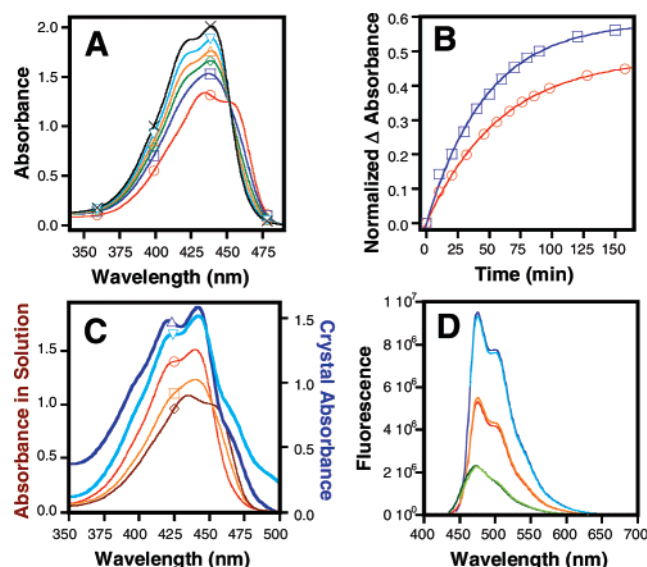
FIGURE 5: Optical properties of Cerulean GFP in solution and in the crystalline state. (A) Time-dependent Cerulean absorbance spectra observed at pH 5.0 upon introduction of a pH jump from 8.0 to 5.0. $t = 0$ red ($\bigcirc$), 20 min blue ($\square$), 40 min green ($\diamondsuit$), 60 min orange ($\triangle$), 90 min light blue ($\triangledown$), 150 min black (X). (B) Single-exponential kinetic fit of the change in optical density at 440 nm (normalized by the optical density at the isosbestic point at 452 nm) upon pH jump from 8.0 to 5.0 (blue $\square$), and from 8.0 to 6.0 (red $\bigcirc$). (C) Solution absorbance spectra of Cerulean at approximately 1.0 mg/mL, normalized by absorbance at 452 nm and extensively equilibrated at pH 5.0 (red $\bigcirc$), pH 6.0 (orange $\square$), and pH 7.0 (brown $\diamondsuit$), compared to raw spectra collected on single crystals at pH 5.0, as shown in bold blue ($\triangledown$) and dark blue ($\triangle$). (D) Fluorescence emission intensity at pH 8.0 (dark blue: $\lambda_{ex} = 435$ nm, light blue: $\lambda_{ex} = 454$ nm) and at pH 5.0. The pH 5.0 data were collected 5 min (red: $\lambda_{ex} = 424$ nm, orange: $\lambda_{ex} = 439$ nm) and 80 min (dark green: $\lambda_{ex} = 424$ nm, light green: $\lambda_{ex} = 439$ nm) after a pH jump from 8.0 to 5.0.

Protein denaturation is not likely a contributing factor, since denaturation at pH 2.5 yields a single chromophore absorbance band centered on 415 nm (data not shown). Similar optical changes are also observed upon introduction of a pH jump from pH 8.0 to 6.0, but not from pH 8.0 to 7.0. The time course of the observed absorbance change fits well to a unimolecular, monoexponential kinetic process (Figure 5B), and the extracted time constants for the jump to pH 5.0 ($\tau_{8\rightarrow5}$), to pH 6.0 ($\tau_{8\rightarrow6}$), and the reverse process ($\tau_{5\rightarrow8}$), are 47 (±2) min, 59 (±2) min, and 63 (±1), respectively. The similar values of about 1 h suggest a slow conformational adjustment with minimal hysteresis in response to a local rearrangement of charges.

Single-crystal absorbance spectra were collected on a micro-spectrophotometer after sample freezing at 100 K (Figure 5C, bold blue scans). For this experiment, thin platelike crystals were chosen that were grown in the same mother liquor (pH 5.0 acetate) as the crystals utilized for X-ray data collection (to date, we have been unable to obtain crystals at higher pH, and soaking experiments have resulted in the dissolution of crystals). The absorbance spectra collected on crystals exhibit the same double-humped shape as seen in solution spectra, confirming that both electronic transitions occur from a single protein conformation, as defined by the resolution limit of the X-ray structure. Importantly, the spectra are essentially identical to the pH 5.0 solution spectra collected after equilibration (Figure 5C,

red curve), with slightly improved separation of the two overlapping bands, likely due to crystal freezing.

Fluorescence emission scans were collected on solution samples at pH 8.0 and 5.0, using four different excitation wavelengths to sample the absorbance bands (Figure 5D). The emission band shape appears nearly invariant: It is not affected by pH, excitation wavelength, or incubation time, suggesting that emission occurs from a common excited state. However, fluorescence intensity at pH 5.0 decreases about 4-fold over a time period of 80 min. A similarly dramatic loss of fluorescence has previously been described for the blue fluorescent protein bearing the Y66H substitution (*39*). The histidine-derived chromophore of this variant is reminiscent of the Cerulean tryptophan-derived chromophore, since the chromophore itself does not undergo changes in protonation equilibrium over the physiological pH range.

DISCUSSION

Implications of the cis Conformation of the Cerulean Chromophore. The cis conformation of the tryptophan-derived chromophore (rotated around the Trp $C_\beta-C_\gamma$ bond, Figure 3) has not previously been observed in any cyan fluorescent GFPs bearing the Y66W mutation (*23*) and was not predicted by molecular dynamics simulations (*40*). However, the crystallographically determined electron density of Cerulean GFP indicates that the entire chromophore population adopts a cis-coplanar conformation (Figure 2), which appears to be stabilized by an extensive network of hydrogen bonds (Figure 4). Although replacement of His148 with an aspartic acid was predicted to preferentially stabilize a structural arrangement in which the engineered carboxyl group would be solvent-exposed (*20*), the crystallographic features of Cerulean determined at pH 5.0 unequivocally place the aspartic acid side chain into a buried position. Its two carboxy oxygen atoms directly interact with the chromophore indole nitrogen via a bifurcated hydrogen bond.

The extensive hydrogen-bonded network connecting the indole nitrogen with the proteinaceous environment encompasses a total of four protein side chains, three main chain groups, and at least five ordered water molecules (Figure 4). This arrangement of extended noncovalent contacts suggests a cagelike structure of optimized hydrogen bonding connections that envelopes the most hydrophobic part of the chromophore, the six-membered ring of the indole moiety. These interactions completely satisfy the hydrogen-bonding potential of the H148D carboxyl and the Thr203 hydroxyl groups, and may modulate the ionization equilibrium of the engineered aspartic acid by substantially raising its pK_a value.

The high degree of chromophore planarity observed in Cerulean is consistent with a favorable interaction between the two ring systems, suggestive of weak hydrogen bonding between an indole aromatic ring carbon and the imidazolinone nitrogen (Figure 4). However, based on the resolution limit of the X-ray structure, hydrogen bonding cannot be distinguished from van der Waals interactions. Therefore, any intrinsic energetic preference for the cis conformation that may exist in tryptophan-based chromophores is difficult to evaluate by inspection of the X-ray structure alone. Nonetheless, for tyrosine-based model chromophores, an energetic preference for the cis conformation (rotated around the Tyr $C_\alpha-C_\beta$ bond) has been demonstrated (*41*). These

types of chromophores tend to adopt twisted or bent structures in the trans conformation (*33*, *42*), in accord with an intra-chromophore steric clash rather than hydrogen bonding as may exist in Cerulean.

Structural Consequences of the H148D Substitution. The structural data suggest that the cis conformation is largely favored in Cerulean due to the substitution of His148 with an aspartic acid, a replacement of a large side chain with a smaller one that allows for repacking of the protein matrix around the isomerized chromophore (Figure 3). Thus, the H148D substitution likely triggers a switch in indole hydrogen-bonding partner from Ser205 (ECFP, trans) to H148D (Cerulean, cis), at least in acidic media. Clearly, the buried aspartic acid is able to specifically stabilize the observed chromophore conformation via hydrogen bonding interactions (Figure 3, red model). We speculate that under basic conditions, the H148D carboxyl may ionize, and its side chain may rotate around $C_\alpha-C_\beta$ to adopt a position similar to that of His148 in ECFP B′ (Figure 3, cyan model). Although we were unable to grow diffraction-quality Cerulean crystals at elevated pH, this scenario appears reasonable, since the Cerulean main chain trace closely follows that of ECFP B′ along residues 144−148 (Figure 1). Careful inspection of the ECFP models suggests that the His148 imidazole group is solvent-exposed in both A′ and B′ conformations; hence, a rotamer of the Cerulean H148D similar to B′ His148 would minimize main chain rearrangements while allowing for solvation of the carboxylate anion. His148 in ECFP B′ adopts a conformation that avoids a steric clash with the bulky trans chromophore (Figure 3). If the engineered Cerulean aspartate were to adopt a similar position upon ionization, the interior cavity created would be large enough to accommodate chromophore isomerization from cis to trans. This process may be energetically favorable, as the disrupted hydrogen bond between the indole nitrogen and H148D would be replaced by hydrogen bonding to Ser205 in the trans conformation.

Is the Low-pH Hypsochromic Shift Related to Chromophore Isomerization? The two overlapping absorption bands observed at 435/454 nm (Figure 5C) (*20*) are shifted to 424/439 nm when the pH is dropped from 7 to 5 (Figure 5A). This shift likely originates from a local rearrangement of charges around the chromophore upon acidification of the medium. However, the position and shape of the fluorescence emission band are not significantly modulated by pH (Figure 5D), supporting a model in which the excited-state from which emission occurs is essentially unaffected by changes in protonation equilibria (Figure 5D). The slow optical response to acidification of the medium suggests a high-energy barrier for conversion to the blue form. Although we are currently unable to distinguish specific mechanisms, the structural arguments presented above are supported by the observed changes in absorbance. We speculate that H148D is anionic at pH values at or above 7, where it is solvated via side chain rotamer adjustments. The titration midpoint of its carboxyl group may lie around pH 6 (Figure 5C), and in more acidic media, protonation of this group may trigger rotation of the side chain toward the protein's interior, where it forms a bifurcated hydrogen bond with the indole ring of the chromophore (Figure 4). This three-center hydrogen bonding interaction may stabilize the chromophore's ground

state, providing a rationale for the blue-shifted absorbance spectra (Figure 5A).

As suggested above, it is possible that acidification is accompanied by slow chromophore ionization from trans to cis. Although the observed time constant of about 1 h is consistent with reported trans-to-cis fluorescence recovery times for tyrosine-based chromophores (*42*), the available data cannot be utilized to determine whether electrostatic changes around the Cerulean chromophore induce isomerization. The essentially identical spectra of high-pH Cerulean and ECFP are difficult to interpret in terms of isomer populations. To underscore this point, absorbance spectra of model GFP-like chromophores have been shown to exhibit little intrinsic sensitivity toward isomerization (*43*). Therefore, resolution of this issue will need to await the availability of high-pH Cerulean crystals forms.

Multiple Main Chain Conformations Have Been Eliminated in Cerulean. In the ECFP X-ray structure published several years ago, residues 144−148 were modeled in two different structural arrangements (Figure 1) (*23*). In addition, a dynamic exchange between different ECFP backbone conformations along β-strand #7 (residues 146−158) has been characterized by NMR and molecular dynamics simulations (*40*, *44*, *45*). However, the Cerulean X-ray structure does not provide any evidence of multiple main chain conformations, yet its chromophore retains two spectral bands, as is observed in ECFP (*20*) (see below). It appears that the Y145A substitution may largely be responsible for locking residues 145−148 into a conformation similar to ECFP B′ (Figure 1). Unexpectedly, the Cerulean Y145A side chain points out toward bulk solvent, as does the tyrosine phenol in ECFP B′. Therefore, Y145A appears to improve protein stability by eliminating a large, hydrophobic group that may otherwise be solvent-exposed. This arrangement allows for close packing of Ile146 against the most hydrophobic part of the chromophore, possibly contributing toward stabilization of the cis form (Figure 3).

However, the Cerulean α-carbon positions follow those of ECFP B′ only along residues 144−148, then they diverge from both A′ and B′ along residues 149−151 (Figure 1). In this range, the main chain trace is pushed out from the protein's interior, likely to accommodate the H148D side chain which is completely buried. A similar explanation has been proposed for pH-dependent conformational changes observed in deGFPs, where movements along residues 143−150 were correlated with internal hydrogen bonding rearrangements in response to chromophore ionization (*32*). Although the indole-based chromophore of Cerulean cannot ionize over the physiological pH range, the particular main chain conformation observed in this variant is likely stabilized by interior packing rearrangements facilitated by the H148D and Y145A substitutions (Figure 1). The data presented here support a planar chromophore π-system that is held in place by a single protein conformation, likely responsible for the reduced heterogeneity of the Cerulean excited-state lifetime (*20*).

Origin of the Double-Humped Band Shape. Surprisingly, the absorbance spectrum of Cerulean is very similar to that of ECFP, with two overlapping absorption bands occurring at 435 and 454 nm at physiological pH values (*20*) (Figure 5A). The double-humped band shape persists in frozen protein crystals, as single-crystal absorbance spectra display

very similar maxima (Figure 5C). For this reason, the two peaks are not likely the result of ground-state structural heterogeneity along the "flipping" region of residues 145–149, as has been suggested previously (*23*). This idea is further supported by quantum mechanical calculations on the ECFP chromophore, where spectra calculated for the two main chain conformations A′ and B′ were judged to be essentially identical (*40*).

In combination, our data suggest that the broad doublet may have its origin in the chemical nature of the conjugated chromophore π system, in analogy to the photophysical properties that have been described for indole groups. In tryptophan, two overlapping $\pi \rightarrow \pi^*$ electronic transitions, ^1L_a and ^1L_b, contribute to the UV absorption spectrum (*46*), with one transition leading to a maximum at $\sim$270 nm, and the other to a doublet peak at 280 and 290 nm. One may speculate that a similar combination of transitions occurs upon conjugation of the indole moiety with the electron-poor imidazolinone ring of the ECFP chromophore. According to this argument, the broad doublet does not arise as a function of structural heterogeneity, but instead may represent an intrinsic property of indole photophysics. In support of this interpretation, recent quantum mechanical calculations on ECFP have predicted that two active absorption bands may coexist in some indole-derived chromophore geometries due to mixing of $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions (*40*).

CONCLUSIONS

The most significant structural difference of Cerulean to ECFP lies in the unusual cis conformation of the tryptophan-derived chromophore and the cagelike network of hydrogen bonds surrounding it. It appears that the single protein main chain conformation, arranged to accommodate the indole-bearing chromophore by fine-tuning hydrogen bonding and interior packing interactions, plays a crucial role in the enhancement of the Cerulean optical properties. The structural observations argue in favor of a rigid, planar chromophore with excited-state decay kinetics comprising one primary component (>95%) over a broad pH range. In contrast, the observation of multiple lifetimes appears to be more closely associated with multiple protein conformations that likely coexist in a dynamic equilibrium around the trans chromophore of ECFP (Figure 1) (*20*).

In spite of several unforeseen structural characteristics concerning the designed mutations H148D and Y145A, rational optimization by genetic engineering has led to highly desirable protein features (*20*). The high quantum yield (0.62) and homogeneous excited-state lifetime are essential parameters in achieving high FRET efficiencies for in vivo measurements of protein−protein interactions.

ACKNOWLEDGMENT

REFERENCES

1. Frokjaer, S., and Otzen, D. E. (2005) Protein drug stability: A formulation challenge. *Nat. Rev. Drug Discovery 4*, 298−306.
2. Bornscheuer, U. T., and Pohl, M. (2001) Improved biocatalysts by directed evolution and rational protein design. *Curr. Opin. Chem. Biol. 5*, 137−143.
3. Tsien, R. Y. (2005) Building and breeding molecules that spy on cells and tumors. *FEBS Lett. 579*, 927−932.
4. Verkhusha, V., and Lukyanov, K. A. (2004) The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nat. Biotechnol. 22*, 289−296.
5. Henderson, J. N., and Remington, S. J. (2006) The kindling fluorescent protein: A transient photoswitchable marker. *Physiology 21*, 162−170.
6. Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002) A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A. 99*, 7877−7882.
7. Ai, H.-W., Henderson, J. N., Remington, S. J., and Campbell, R. E. (2006) Directed evolution of a monomeric, bright, and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescent imaging. *Biochem. J. 400*, 531−540.
8. Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., and Tsien, R. Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol. 22*, 1567−1572.
9. Patterson, G. H., Piston, D. W., and Barisas, B. G. (2000) Förster distances between green fluorescent protein pairs. *Anal. Biochem. 284*, 438−440.
10. Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996) Crystal structure of the Aequorea victoria Green Fluorescent Protein. *Science 273*, 1392−1395.
11. Wachter, R. M., Elsliger, M.-A., Kallio, K., Hanson, G. T., and Remington, S. J. (1998) Structural basis of spectral shifts in the yellow-emission variants of Green Fluorescent Protein. *Structure 6*, 1267−1277.
12. Griesbeck, O., Baird, G. S., Campbell, R. E., Zacharias, D. A., and Tsien, R. Y. (2001) Reducing the environmental sensitivity of yellow fluorescent protein. *J. Biol. Chem. 276*, 29188−29194.
13. Remington, S. J. (2006) Fluorescent proteins: Maturation, photochemistry and photophysics. *Curr. Opin. Struct. Biol. 16*, 1−8.
14. Wachter, R. M. (2006) Chromogenic cross-link formation in green fluorescent protein. *Acc. Chem. Res. 40*, 120−127.
15. Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., and Lukyanov, S. A. (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol. 17*, 969−973.
16. Henderson, J. N., and Remington, S. J. (2005) Crystal structures and mutational analysis of amFP486, a cyan fluorescent protein from Anemonia majano. *Proc. Natl. Acad. Sci. U.S.A. 102*, 12712−12717.
17. Ai, H.-w., Shaner, N. C., Cheng, Z., Tsien, R. Y., and Campbell, R. E. (2007) Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins. *Biochemistry 46*, 5904−5910.
18. Heim, R., Prasher, D. C., and Tsien, R. Y. (1994) Wavelength mutations and posttranslational autoxidation of Green Fluorescent Protein. *Proc. Natl. Acad. Sci. U.S.A. 91*, 12501−12504.
19. Tsien, R. Y. (1998) The Green Fluorescent Protein. *Ann. Rev. Biochem. 67*, 509−544.
20. Rizzo, M. A., Springer, G. H., Granada, B., and Piston, D. W. (2004) An improved cyan fluorescent protein variant useful for FRET. *Nat. Biotechnol. 22*, 445−449.
21. Kremers, G.-J., Goedhart, J., van Munster, E. B., and Gadella, T. W. J. (2006) Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Forster radius. *Biochemistry 45*, 6570−6580.
22. Nguyen, A. W., and Daugherty, P. S. (2005) Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol. 23*, 355−360.
23. Bae, J. H., Rubini, M., Jung, G., Wiegand, G., Seifert, M. H. J., Azim, M. K., Kim, J.-S., Zumbusch, A., Holak, T. A., Moroder, L., Huber, R., and Budisa, N. (2003) Expansion of the genetic code enables design of a novel "Gold" class of green fluorescent proteins. *J. Mol. Biol. 328*, 1071−1081.
24. Cubitt, A. B., Woollenweber, L. A., and Heim, R. (1999) Understanding structure-function relationships in the Aequorea victoria green fluorescent protein. *Methods Cell Biol. 58*, 19−30.
25. Rizzo, M. A., Springer, G. H., Segawa, K., Zipfel, W. R., and Piston, D. W. (2006) Optimization of pairings and detection conditions for measurement of FRET between cyan and yellow fluorescent proteins. *Microsc. Microanal. 12*, 238−254.

26. Rizzo, M. A., and Piston, D. W. (2005) High-contrast imaging of fluorescent protein FRET by fluorescence polarization microscopy. *Biophys. J. 88*, L14−16.

27. Leslie, A. G. (1999) Integration of macromolecular diffraction data. *Acta Crystallogr. D55*, 1696−1702.

28. Collaborative Computational Project No. 4. (1994) The CCP4 Suite: Programs for protein crystallography. *Acta Crystallogr. D50*, 760−763.

29. Vagin, A., and Teplyakov, A. (1997) An automated program for molecular replacement. *J. Appl. Crystallogr. 30*, 1022−1025.

30. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D53*, 240−255.

31. McRee, D. E. (1999) XtalView/Xfit−A versatile program for manipulating atomic coordinates, and electron density. *J. Struct. Biol. 125*, 156−165.

32. Hanson, G. T., McAnaney, T. B., Park, E. S., Rendell, M. E. P., Yarbrough, D. K., Chu, S., Xi, L., Boxer, S. G., Montrose, M. H., and Remington, S. J. (2002) Green fluorescent protein variants as ratiometric dual emission pH sensors. 1. Structural characterization and preliminary application. *Biochemistry 41*, 15477−15488.

33. Quillin, M. L., Anstrom, D. M., Shu, X., O'Leary, S., Kallio, K., Chudakov, D. M., and Remington, S. J. (2005) Kindling fluorescent protein from Anemonia sulcata: Dark-state structure at 1.38 A resolution. *Biochemistry 44*, 5774−5787.

34. Desiraju, G. R., and Steiner, T. (1999) *The Weak Hydrogen Bond*, Oxford University Press, Inc., New York.

35. Taylor, R., and Kennard, O. (1982) Crystallographic evidence for the existence of C−H···O, C−H···N, and C−H···Cl hydrogen bonds. *J. Am. Chem. Soc. 104*, 5063−5070.

36. Gao, E.-Q. (2005) 1,4-Diphenyl-1,4-di-4-pyridyl-2,3-diaza-1,3-butadiene: weak C−H···N and C−H···π hydrogen bonds. *Acta Crystallgr. C61*, o110−o111.

37. Zimmer, M. (2002) Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior. *Chem. Rev. 102*, 759−781.

38. Preissner, R., Egner, U., and Saenger, W. (1991) Occurrence of bifurcated three-center hydrogen bonds in proteins. *FEBS Lett. 288*, 192−196.

39. Wachter, R. M., Brett, A. K., Heim, R., Kallio, K., Tsien, R. Y., Boxer, S. G., and Remington, S. J. (1997) Crystal structure and photodynamic behavior of the blue emission variant Y66H/Y145F of Green Fluorescent Protein. *Biochemistry 36*, 9759−9765.

40. Demachy, I., Ridard, J., Laguitton-Pasquier, H., Durnerin, E., Vallverdu, G., Archirel, P., and Levy, B. (2005) Cyan fluorescent protein: Molecular dynamics, simulations, and electronic absorption spectrum. *J. Phys. Chem. B 109*, 24121−24133.

41. He, X., Bell, A. F., and Tonge, P. J. (2003) Ground state isomerization of a model green fluorescent protein chromophore. *FEBS Lett. 549*, 35−38.

42. Henderson, J. N., Ai, H.-w., Campbell, R. E., and Remington, S. J. (2007) Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proc. Natl. Acad. Sci. U.S.A. 104*, 6672−6677.

43. He, X., Bell, A. F., and Tonge, P. J. (2002) Isotopic labeling and normal-mode analysis of a model green fluorescent protein chromophore. *J. Phys. Chem. B 106*, 6056−6066.

44. Seifert, M. H. J., Ksiazek, D., Azim, M. K., Smialowski, P., Budisa, N., and Holak, T. A. (2002) Slow exchange in the chromophore of a green fluorescent protein variant. *J. Am. Chem. Soc. 124*, 7932−7942.

45. Seifert, M. H. J., Georgescu, J., Ksiazek, D., Smialowski, P., Rehm, T., Steipe, B., and Holak, T. A. (2003) Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution. *Biochemistry 42*, 2500−2512.

46. Creed, D. (1984) The photophysics and photochemistry of the near-UV absorbing amino acids-I. Tryptophan and its simple derivatives. *Photochem. Photobiol. 39*, 537−562.